

# Decision-Driven Credal Information Fusion

AQM Sazzad Sayyed<sup>†</sup>, Michele Caprio<sup>\*</sup>, Nathaniel D. Bastian<sup>‡</sup> and Francesco Restuccia<sup>†</sup>

<sup>†</sup>Northeastern University, USA   <sup>\*</sup>The University of Manchester, UK   <sup>‡</sup>United States Military Academy, USA

**Abstract**—Modern safety-critical systems rely on multiple sensors and must fuse outputs from multiple processing streams under uncertainty. Credal sets offer a natural way to represent this uncertainty and provide a geometric structure for uncertainty-aware multi-source fusion. While prior work often treats the fusion rule as a heuristic design choice, we show that the optimal fusion rule is theoretically dictated by the selected aleatoric uncertainty (AU) functional and its order-theoretic behavior. We study two canonical credal fusion operators, i.e., conjunctive fusion, defined as the intersection of credal sets, and disjunctive fusion, defined as the core induced by the pointwise minimum of their lower probabilities. We prove that the choice between them is deterministic, i.e., if the chosen AU functional is inclusion-monotone (AU increases as the credal set grows), then conjunctive fusion minimizes uncertainty; if it is inclusion-antitone (AU decreases as the credal set grows), then disjunctive fusion is optimal. Thus, operator selection is not empirical or ad hoc, but follows directly from the monotonicity properties of the downstream objective. For finite outcome spaces, we further provide a polyhedral characterization of the disjunctive fusion core and a dual feasibility criterion ensuring coherence. We validate these predictions in a multi-sensor image classification setting under distribution shift on CIFAR10, CIFAR100, and SVHN. Using ResNet architectures, we find that AU-guided fusion improves calibration and stabilizes performance under shift relative to fixed fusion strategies. On CIFAR10, predictive accuracy improves by up to 24%, and on SVHN by up to 16%. Calibration error is reduced by up to 68.64% on CIFAR10 and 20% on SVHN.

## I. INTRODUCTION

Modern decision-making systems frequently aggregate multiple sources of information—sensors, models, or agents—to improve accuracy, robustness, and actionability [1], [2]. However, when sources are statistically dependent (e.g., share training data, modeling assumptions, or latent noise), naïve fusion can lead to systematic overconfidence [3]. A central limitation of standard probabilistic pipelines is their reliance on a *single* predictive distribution per source. While suitable for modeling irreducible noise – i.e., aleatoric uncertainty (AU) – this representation obscures *epistemic uncertainty (EU)* arising from limited data, model misspecification, or distribution shift [4], [5]. Reducing epistemic variability into a singleton can therefore bias downstream decisions.

Credal sets provide a principled alternative. A credal set is a nonempty, closed, convex set of probability measures representing a family of plausible predictive distributions [6], [7]. Such representations explicitly encode epistemic variation and are increasingly realizable in modern learning systems, including distributionally robust methods and models with prior or likelihood ambiguity [8], [9]. The central question we address is: *how should credal information be fused?* Fusion operators for imprecise probabilities and belief models—most prominently conjunctive and disjunctive rules—have been extensively studied [10]–[12]. Yet these operators are typically

presented as modeling choices rather than decision-dependent constructions. In parallel, a mature literature on uncertainty measures for credal sets provides tools for comparing fused outcomes [13]. What is missing is a criterion linking the choice of fusion operator to the downstream decision.

We resolve this gap by turning fusion operator selection into a consequence of the structure and characteristics of the credal sets and the AU functional. We consider a practical scenario as illustrated in 1 with  $N$  parallel sensor streams observing the same object under distribution shift. For homogeneous sensors, the data streams can share the same inference backbone while for heterogeneous sensors, it will differ. At the output, each stream induces a credal set over class probabilities. We introduce two transparent fusion operators: *conjunctive fusion* (intersection of credal sets) and *disjunctive fusion* (the core of the pointwise minimum of lower probabilities) for our analysis of optimality. A key structural result of our analysis is that the conjunctively fused credal set is always contained in the disjunctively fused credal set, which forms the geometric backbone of the problem.

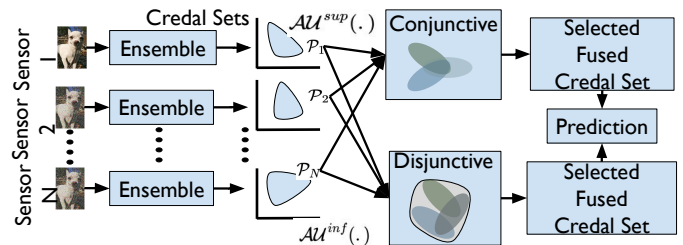


Fig. 1: AU-guided credal fusion: each sensor ensemble produces a credal set; conjunctive and disjunctive fusions are computed; the operator is selected by minimizing a chosen uncertainty functional.

Building on this structure, we show that optimal fusion is determined entirely by the order-theoretic behavior of the chosen aleatoric uncertainty (AU) functional. If the AU functional is inclusion-monotone, conjunctive fusion uniquely minimizes uncertainty. If it is inclusion-antitone, disjunctive fusion is optimal. Under mild conditions, both constructions yield valid credal sets, and for finite outcome spaces we provide an explicit polyhedral characterization of the disjunctive core together with a dual feasibility (avoiding-sure-loss) condition. Fusion rule selection thus becomes a theorem-driven consequence of inclusion geometry.

We empirically validate this inclusion-driven selection principle in a multi-sensor image classification setting. We simulate controlled distribution shifts (clean views vs covariate shift) and evaluate predictive performance using accuracy, neg-

ative log-likelihood, and Brier score under fixed fusion rules versus AU-guided operator selection. Experiments on CIFAR and SVHN with ResNet architectures [14]–[16] demonstrate that AU-guided fusion improves calibration and stabilizes performance under shift relative to fixed strategies. For CIFAR10, in terms of predictive accuracy, optimal AU guided fusion outperforms fixed fusion strategies by up to 24%, while for SVHN, this improvement is up to 16%. Optimal AU based fusion improves calibration error by upto 68.64% for CIFAR10 and 20% for SVHN.

## II. RELATED WORK

Information fusion has been studied across multiple uncertainty calculi—including probability, possibility theory, belief functions, and random sets—with sustained attention to principled combination rules and their behavioral interpretation (e.g., consensus vs. compromise, caution vs. informativeness). A comprehensive synthesis is provided by [17], while [18] advocate a unifying design perspective in which many concrete operators arise from broader structural principles.

Within imprecise probability (IP) theory [6], model combination is typically framed as an aggregation problem: given several sets of probabilities, construct a coherent aggregate reflecting a specified pooling attitude. Foundational approaches include aggregation via coherent lower previsions and combination rules [19], and axiomatic or decision-theoretic treatments of pooling [20]. Related normative questions about combining imprecise beliefs have also been examined in formal epistemology [21]. In parallel, the Dempster–Shafer framework emphasizes conjunctive and disjunctive evidence combination. [11] analyze these rules under dependence and conflict, and [22] provide an engineering-oriented overview of combination principles and their properties.

Selecting among fusion outputs requires principled measures of uncertainty for probability sets. Classical proposals include general information measures for imprecise probabilities [23] and axiomatic treatments of total uncertainty and its decomposition (e.g., nonspecificity and conflict) [24], [25]. In machine learning, credal-set uncertainty has been empirically reviewed in [26], highlighting challenges in separating aleatoric and epistemic components. Geometric measures such as credal volume have been proposed as epistemic indicators [27], [28]. Complementarily, [29] derive uncertainty measures from decompositions of proper scoring rules, providing a scoring-theoretic route to defining aleatoric and epistemic quantities for set-valued predictions.

Because our criterion is formulated as an optimization over fusion outputs, it connects to the literature on evaluating probabilistic forecasts. Proper scoring rules are standard for precise probabilities [30]. For imprecise probabilities, impossibility results and alternative evaluative principles have been studied in [31]. More recent work develops scoring rules for imprecise forecasts [32], [33], analyzes calibration under imprecision [34], and investigates truthful elicitation mechanisms for imprecise beliefs [35].

**Novelty of Our Approach.** Existing research primarily (i) proposes or analyzes fusion operators, or (ii) develops uncertainty measures for credal models. In contrast, we study their interaction explicitly: fusion is performed at the probability-set level, and the choice between conjunctive and disjunctive operators is determined by minimizing an application-driven aleatoric uncertainty functional. This elevates operator selection from a modeling convention to a decision-theoretic optimization principle grounded in set inclusion geometry.

## III. PROBLEM DESCRIPTION

Let  $(\mathcal{X}, \mathcal{F}_{\mathcal{X}})$  and  $(\mathcal{Y}, \mathcal{F}_{\mathcal{Y}})$  be the input and output measurable spaces, respectively, and we denote by  $x$  and  $y$  their respective elements. Consider two procedures that given a training set  $D = \{(x_i, y_i)\}_{i=1}^n$  and a new test input  $x_{n+1}$ , for some  $n \in \mathbb{N}$ , allow us to derive two (nonempty predictive) credal sets on  $\mathcal{Y}$ .<sup>1</sup> Call them  $\mathcal{P}$  and  $\mathcal{Q}$ , respectively. More formally, we assume the existence of two maps  $(D, x_{n+1}) \mapsto \mathcal{P}$  and  $(D, x_{n+1}) \mapsto \mathcal{Q}$ .

**Remark 1.** *Formally, a credal set  $\mathfrak{P}$  is defined as a nonempty convex and closed set of probabilities [36]. Its upper and lower envelopes are defined as  $\bar{P}(A) := \sup_{P \in \mathfrak{P}} P(A)$  and  $\underline{P}(A) := \inf_{P \in \mathfrak{P}} P(A)$ , for all measurable sets  $A$ , respectively. They are conjugate, i.e.  $\bar{P}(A) = 1 - \underline{P}(A^c)$ , for all measurable sets  $A$ . The elements  $P \in \mathfrak{P}$  are “plausible” models; working with the whole sets allows us to hedge against possible distribution misspecification and shift, and to conduct a more robust analysis.*

We call  $\Delta_{\mathcal{Y}}$  the space of (countably additive) probability measures on  $\mathcal{Y}$ , so that  $\mathcal{P}, \mathcal{Q} \subseteq \Delta_{\mathcal{Y}}$ . Let  $\mathcal{F}_{\mathcal{Y}} = \mathcal{B}(\mathcal{Y})$ . We endow  $\Delta_{\mathcal{Y}}$  with the setwise topology, i.e. the coarsest topology making  $\mu \mapsto \mu(A)$  continuous for every  $A \in \mathcal{F}_{\mathcal{Y}}$ . Notice how further restrictions may be imposed on such maps, e.g. Lipschitzianity, and that our following arguments can be effortlessly generalized to more than just two credal-set-inducing procedures. We are interested in formalizing the concept of *information fusion* for credal sets. To do so, we propose two different fusion methods.

### A. Conjunctive Fusion

**Assumption 1.**  $\mathcal{P} \cap \mathcal{Q} \neq \emptyset$ .

Let  $\mathcal{T}_c := \mathcal{P} \cap \mathcal{Q}$ . We prove that  $\mathcal{T}_c$  is a credal set.

*Proof.* Since  $\mathcal{P}, \mathcal{Q} \subseteq \Delta_{\mathcal{Y}}$ , we have  $\mathcal{P} \cap \mathcal{Q} \subseteq \Delta_{\mathcal{Y}}$ . The intersection of closed sets is closed, hence  $\mathcal{P} \cap \mathcal{Q}$  is closed. Moreover, the intersection of convex sets is convex: if  $P, Q \in \mathcal{P} \cap \mathcal{Q}$  and  $\lambda \in [0, 1]$ , then  $\lambda P + (1 - \lambda)Q \in \mathcal{P}$  and  $\lambda P + (1 - \lambda)Q \in \mathcal{Q}$ , so  $\lambda P + (1 - \lambda)Q \in \mathcal{P} \cap \mathcal{Q}$ . By assumption  $\mathcal{P} \cap \mathcal{Q} \neq \emptyset$ ; thus  $\mathcal{P} \cap \mathcal{Q}$  is nonempty, closed and convex, i.e. a credal set.  $\square$

### B. Disjunctive Fusion

Let  $\underline{P}$  and  $\underline{Q}$  be the lower probabilities for  $\mathcal{P}$  and  $\mathcal{Q}$ , respectively. Let  $\underline{R}(A) := \min\{\underline{P}(A), \underline{Q}(A)\}$ ,  $\forall A \in \mathcal{F}_{\mathcal{Y}}$ .

<sup>1</sup>That is, closed (in the setwise topology) and convex sets of probabilities.

**Proposition 1.**  $\underline{R}$  is a lower probability.<sup>2</sup>

*Proof.* The core (set of dominating probabilities) of  $\underline{R}$  is

$$\mathcal{M}(\underline{R}) := \{M \in \Delta_{\mathcal{Y}} : M(A) \geq \underline{R}(A), \quad \forall A \in \mathcal{F}_{\mathcal{Y}}\}.$$

We show that  $\underline{R}$  is the lower envelope of  $\mathcal{M}(\underline{R})$ , i.e.

$$\underline{R}(A) = \inf_{M \in \mathcal{M}(\underline{R})} M(A), \quad \forall A \in \mathcal{F}_{\mathcal{Y}}. \quad (1)$$

**Step 1:  $\mathcal{M}(\underline{R})$  is nonempty (indeed it contains  $\mathcal{P} \cup \mathcal{Q}$ ).** Take any  $P \in \mathcal{P}$ . For every event  $A$ ,

$$\underline{R}(A) = \min\{\underline{P}(A), \underline{Q}(A)\} \leq \underline{P}(A) \leq P(A),$$

where the last inequality holds because  $\underline{P}(A) = \inf_{P' \in \mathcal{P}} P'(A) \leq P(A)$ . Thus  $P \in \mathcal{M}(\underline{R})$ . Hence  $\mathcal{P} \subseteq \mathcal{M}(\underline{R})$ . Similarly,  $\mathcal{Q} \subseteq \mathcal{M}(\underline{R})$ . In turn,  $\mathcal{M}(\underline{R}) \neq \emptyset$ .

**Step 2: The inequality “ $\geq$ ” in (1).** Fix  $A \in \mathcal{F}_{\mathcal{Y}}$ . By definition of  $\mathcal{M}(\underline{R})$ , every  $M \in \mathcal{M}(\underline{R})$  satisfies  $M(A) \geq \underline{R}(A)$ . Therefore

$$\inf_{M \in \mathcal{M}(\underline{R})} M(A) \geq \underline{R}(A).$$

**Step 3: The inequality “ $\leq$ ” in (1).** Fix  $A \in \mathcal{F}_{\mathcal{Y}}$ .

*Case 1:  $\underline{P}(A) \leq \underline{Q}(A)$ .* Then  $\underline{R}(A) = \underline{P}(A)$ . By the definition of infimum, for each  $n \in \mathbb{N}$  there exists  $P_n \in \mathcal{P}$  such that

$$P_n(A) \leq \underline{P}(A) + \frac{1}{n}.$$

By Step 1,  $P_n \in \mathcal{M}(\underline{R})$  for all  $n$ , hence

$$\inf_{M \in \mathcal{M}(\underline{R})} M(A) \leq P_n(A) \leq \underline{P}(A) + \frac{1}{n}.$$

Letting  $n \rightarrow \infty$  yields  $\inf_{M \in \mathcal{M}(\underline{R})} M(A) \leq \underline{P}(A) = \underline{R}(A)$ .

*Case 2:  $\underline{P}(A) > \underline{Q}(A)$ .* Then  $\underline{R}(A) = \underline{Q}(A)$ . Analogously, for each  $n \in \mathbb{N}$  choose  $Q_n \in \mathcal{Q}$  with

$$Q_n(A) \leq \underline{Q}(A) + \frac{1}{n}.$$

Again  $Q_n \in \mathcal{M}(\underline{R})$  by Step 1, so

$$\inf_{M \in \mathcal{M}(\underline{R})} M(A) \leq Q_n(A) \leq \underline{Q}(A) + \frac{1}{n},$$

and letting  $n \rightarrow \infty$  gives  $\inf_{M \in \mathcal{M}(\underline{R})} M(A) \leq \underline{Q}(A) = \underline{R}(A)$ . Combining Steps 2 and 3 proves (1). Since (a)  $\mathcal{M}(\underline{R})$  is a nonempty set of probability measures, (b) it is setwise closed, because  $\mathcal{M}(\underline{R}) = \bigcap_{A \in \mathcal{F}_{\mathcal{Y}}} \{M \in \Delta_{\mathcal{Y}} : M(A) \geq \underline{R}(A)\}$  and each set  $\{M : M(A) \geq \underline{R}(A)\}$  is closed since  $M \mapsto M(A)$  is continuous in the setwise topology, and (c) it is convex, since  $\lambda M_1 + (1 - \lambda)M_2 \geq \lambda \underline{R} + (1 - \lambda)\underline{R} = \underline{R}$ , for all  $\lambda \in [0, 1]$  and all  $M_1, M_2 \in \mathcal{M}(\underline{R})$ , then (1) shows that  $\underline{R}$  is a lower envelope of probabilities, hence a lower probability.  $\square$

Notice that is not needed for the conclusion above: Step 1 already shows  $\mathcal{M}(\underline{R})$  is nonempty because it contains  $\mathcal{P} \cup \mathcal{Q}$ .

We say that the core  $\mathcal{M}(\underline{R}) =: \mathcal{T}_d$  is the credal set resulting from a disjunctive information fusion process. We now characterize its elements.

<sup>2</sup>Following Cerreia-Vioglio et al, 2016 [37], we define a lower probability as the lower envelope of its core.

*C. Core Inclusions and Equality Cases for Disjunctive Fusion*

Let  $\underline{P}$  and  $\underline{Q}$  be lower probabilities on  $(\mathcal{Y}, \mathcal{F}_{\mathcal{Y}})$  and define

$$\underline{R}(A) := \min\{\underline{P}(A), \underline{Q}(A)\}, \quad A \in \mathcal{F}_{\mathcal{Y}}.$$

For any set function  $\underline{S} : \mathcal{F}_{\mathcal{Y}} \rightarrow \mathbb{R}$ , define its core by

$$\mathcal{M}(\underline{S}) := \{M \in \Delta_{\mathcal{Y}} : M(A) \geq \underline{S}(A), \quad \forall A \in \mathcal{F}_{\mathcal{Y}}\}.$$

**Lemma 1** (Order-reversal for cores). *If  $\underline{S}_1(A) \leq \underline{S}_2(A)$  for all  $A \in \mathcal{F}_{\mathcal{Y}}$ , then*

$$\mathcal{M}(\underline{S}_2) \subseteq \mathcal{M}(\underline{S}_1).$$

*Proof.* If  $M \in \mathcal{M}(\underline{S}_2)$ , then  $M(A) \geq \underline{S}_2(A) \geq \underline{S}_1(A)$  for all  $A$ , so  $M \in \mathcal{M}(\underline{S}_1)$ .  $\square$

**Corollary 1** (Inclusions for disjunctive fusion). *We have*

$$\mathcal{M}(\underline{P}) \subseteq \mathcal{M}(\underline{R}), \quad \mathcal{M}(\underline{Q}) \subseteq \mathcal{M}(\underline{R}),$$

and hence

$$\text{co}(\mathcal{M}(\underline{P}) \cup \mathcal{M}(\underline{Q})) \subseteq \mathcal{M}(\underline{R}).$$

*Proof.* Since  $\underline{R} \leq \underline{P}$  and  $\underline{R} \leq \underline{Q}$  pointwise, the first two inclusions follow from Lemma 1. Convexity of  $\mathcal{M}(\underline{R})$  (proved in Proposition 1) then yields  $\text{co}(\mathcal{M}(\underline{P}) \cup \mathcal{M}(\underline{Q})) \subseteq \mathcal{M}(\underline{R})$ .  $\square$

**Proposition 2** (When does  $\mathcal{M}(\underline{R})$  collapse to one input core?). *The following are equivalent:*

- 1)  $\underline{P}(A) \leq \underline{Q}(A)$  for all  $A \in \mathcal{F}_{\mathcal{Y}}$ ;
- 2)  $\underline{R} = \underline{P}$  (pointwise);
- 3)  $\mathcal{M}(\underline{R}) = \mathcal{M}(\underline{P})$ .

Similarly,  $\underline{Q} \leq \underline{P}$  pointwise iff  $\underline{R} = \underline{Q}$  iff  $\mathcal{M}(\underline{R}) = \mathcal{M}(\underline{Q})$ .

*Proof.* The equivalence (1)  $\Leftrightarrow$  (2) is immediate from  $\underline{R} = \min\{\underline{P}, \underline{Q}\}$ . If  $\underline{R} = \underline{P}$  then clearly  $\mathcal{M}(\underline{R}) = \mathcal{M}(\underline{P})$ , giving (2)  $\Rightarrow$  (3). Conversely, if  $\mathcal{M}(\underline{R}) = \mathcal{M}(\underline{P})$ , then taking lower envelopes over the (common) core gives

$$\underline{R}(A) = \inf_{M \in \mathcal{M}(\underline{R})} M(A) = \inf_{M \in \mathcal{M}(\underline{P})} M(A) = \underline{P}(A),$$

so (3)  $\Rightarrow$  (2). (The last equality holds because  $\underline{P}$  is a lower probability, hence equals the lower envelope of its own core.)  $\square$

**Proposition 3** (Intersection characterization).

$$\mathcal{M}(\underline{R}) = \mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q}) \quad \Leftrightarrow \quad \underline{P} = \underline{Q} \text{ pointwise on } \mathcal{F}_{\mathcal{Y}}.$$

*Proof.* If  $\underline{P} = \underline{Q}$  then  $\underline{R} = \underline{P} = \underline{Q}$  and the equality is obvious.

Conversely, assume  $\mathcal{M}(\underline{R}) = \mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q})$ . Taking lower envelopes yields, for all  $A \in \mathcal{F}_{\mathcal{Y}}$ ,

$$\begin{aligned} \underline{R}(A) &= \inf_{M \in \mathcal{M}(\underline{R})} M(A) = \inf_{M \in \mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q})} M(A) \\ &\geq \max\left\{ \inf_{M \in \mathcal{M}(\underline{P})} M(A), \inf_{M \in \mathcal{M}(\underline{Q})} M(A) \right\} \\ &= \max\{\underline{P}(A), \underline{Q}(A)\}. \end{aligned}$$

Since we always have that  $\underline{R}(A) = \min\{\underline{P}(A), \underline{Q}(A)\}$ , we obtain

$$\min\{\underline{P}(A), \underline{Q}(A)\} \geq \max\{\underline{P}(A), \underline{Q}(A)\},$$

for all  $A$ , hence  $\underline{P}(A) = \underline{Q}(A)$  for all  $A$ .  $\square$

**Proposition 4** (A useful sufficient-and-necessary condition in the finite case). *Assume  $\mathcal{Y}$  is finite, so  $\mathcal{M}(\underline{R})$  is a polytope (see Proposition 5 below). Let  $\text{ext}(\mathcal{M}(\underline{R}))$  denote its set of vertices. Then*

$$\begin{aligned} \mathcal{M}(\underline{R}) &= \text{co}(\mathcal{M}(\underline{P}) \cup \mathcal{M}(\underline{Q})) \\ &\iff \text{ext}(\mathcal{M}(\underline{R})) \subseteq \mathcal{M}(\underline{P}) \cup \mathcal{M}(\underline{Q}). \end{aligned} \quad (2)$$

*Proof.* If  $\mathcal{M}(\underline{R}) = \text{co}(\mathcal{M}(\underline{P}) \cup \mathcal{M}(\underline{Q}))$  then every extreme point of the left-hand side is an extreme point of the right-hand side. Since the right-hand side is the convex hull of  $\mathcal{M}(\underline{P}) \cup \mathcal{M}(\underline{Q})$ , its extreme points must lie in  $\mathcal{M}(\underline{P}) \cup \mathcal{M}(\underline{Q})$ . This gives the forward implication.

For the reverse implication, note that in the finite case  $\mathcal{M}(\underline{R})$  is a nonempty compact convex polytope, hence  $\mathcal{M}(\underline{R}) = \text{co}(\text{ext}(\mathcal{M}(\underline{R})))$ . If  $\text{ext}(\mathcal{M}(\underline{R})) \subseteq \mathcal{M}(\underline{P}) \cup \mathcal{M}(\underline{Q})$ , then

$$\mathcal{M}(\underline{R}) = \text{co}(\text{ext}(\mathcal{M}(\underline{R}))) \subseteq \text{co}(\mathcal{M}(\underline{P}) \cup \mathcal{M}(\underline{Q})) \subseteq \mathcal{M}(\underline{R}),$$

where the last inclusion is Corollary 1.  $\square$

#### D. Finite $\mathcal{Y}$ : Polyhedral Description and Feasibility Criteria

Assume  $\mathcal{Y} = \{1, \dots, n\}$  and  $\mathcal{F}_{\mathcal{Y}} = 2^{\mathcal{Y}}$ . Identify a probability  $M$  with its mass vector  $m = (m_1, \dots, m_n) \in \mathbb{R}^n$  where  $m_i = M(\{i\})$ , so  $m_i \geq 0$  and  $\sum_{i=1}^n m_i = 1$ . For  $A \subseteq \mathcal{Y}$  write  $m(A) := \sum_{i \in A} m_i$ .

**Proposition 5** (Core as a polytope). *For any set function  $\underline{S} : 2^{\mathcal{Y}} \rightarrow \mathbb{R}$ , the core is the (possibly empty) polyhedron*

$$\begin{aligned} \mathcal{M}(\underline{S}) &= \\ &\left\{ m \in \mathbb{R}^n : m_i \geq 0, \sum_{i=1}^n m_i = 1, m(A) \geq \underline{S}(A), \quad \forall A \subseteq \mathcal{Y} \right\}. \end{aligned} \quad (3)$$

*In particular, if  $\mathcal{M}(\underline{S}) \neq \emptyset$  it is a compact convex polytope.*

*Proof.* This is a direct rewriting of the defining inequalities  $M(A) \geq \underline{S}(A)$  for all  $A$ , together with the simplex constraints. The set is an intersection of finitely many closed half-spaces (one per  $A \subseteq \mathcal{Y}$ ) with the simplex, hence closed and bounded; convexity is immediate.  $\square$

**Proposition 6** (Dual (avoiding-sure-loss) feasibility criterion). *Let  $\underline{S} : 2^{\mathcal{Y}} \rightarrow \mathbb{R}$ . Then  $\mathcal{M}(\underline{S}) \neq \emptyset$  if and only if*

$$\begin{aligned} \sum_{k=1}^m \lambda_k \underline{S}(A_k) &\leq \max_{i \in \mathcal{Y}} \sum_{k=1}^m \lambda_k \mathbf{1}_{A_k}(i), \\ &\text{for all } m \in \mathbb{N}, \lambda_k \geq 0, A_k \subseteq \mathcal{Y}. \end{aligned} \quad (4)$$

*Proof.* ( $\Rightarrow$ ) If  $m \in \mathcal{M}(\underline{S})$ , then for any  $\lambda_k \geq 0$  and  $A_k$ ,

$$\sum_{k=1}^m \lambda_k \underline{S}(A_k) \leq \sum_{k=1}^m \lambda_k m(A_k) \quad (5)$$

$$= \sum_{i \in \mathcal{Y}} m_i \left( \sum_{k=1}^m \lambda_k \mathbf{1}_{A_k}(i) \right) \quad (6)$$

$$\leq \max_{i \in \mathcal{Y}} \sum_{k=1}^m \lambda_k \mathbf{1}_{A_k}(i), \quad (7)$$

because the  $m_i$ 's are nonnegative coefficients summing up to 1.

( $\Leftarrow$ ) Condition (4) is precisely the dual (separating hyperplane/Farkas lemma) condition for feasibility of the linear system defining  $\mathcal{M}(\underline{S})$  in Proposition 5. Equivalently, if the system were infeasible, there would exist nonnegative multipliers  $\lambda_k$  producing a strict violation of (4).  $\square$

Notice that for  $\underline{R} = \min\{\underline{P}, \underline{Q}\}$  with  $\underline{P}, \underline{Q}$  lower probabilities, one already knows  $\mathcal{M}(\underline{R}) \neq \emptyset$  because  $\mathcal{M}(\underline{P}) \cup \mathcal{M}(\underline{Q}) \subseteq \mathcal{M}(\underline{R})$  (Corollary 1). In the finite case, this implies (4) holds automatically for  $\underline{S} = \underline{R}$ .

## IV. PROPOSED METHOD

We aim at finding which, between conjunctive and disjunctive fusion, is associated with the least induced aleatoric uncertainty (AU). This is because we posit that AU captures the performance of a downstream statistical inference task. Our optimization problem, then, becomes

$$\min_{j \in \{c, d\}} AU(\mathcal{T}_j),$$

where  $AU(\cdot)$  is a measure of aleatoric uncertainty for a generic credal set. If the argmin is not a singleton, we select the fusion procedure according to a splitting mechanism defined later.

**Theorem 2** (Existence (and characterization) of AU-minimizing fusion). *Let  $(\mathcal{Y}, \mathcal{F}_{\mathcal{Y}})$  be as in Section III and endow  $\Delta_{\mathcal{Y}}$  with the setwise topology. Let  $\mathcal{P}, \mathcal{Q} \subseteq \Delta_{\mathcal{Y}}$  be credal sets (nonempty, convex, setwise closed) and assume*

$$\mathcal{P} \cap \mathcal{Q} \neq \emptyset. \quad (8)$$

*Define the fused credal sets*

$$\mathcal{T}_c := \mathcal{P} \cap \mathcal{Q}, \quad \mathcal{T}_d := \mathcal{M}(\underline{R}),$$

where  $\underline{R} = \min\{\underline{P}, \underline{Q}\}$  is as in Section III-B and  $\mathcal{M}(\underline{R}) = \{M \in \Delta_{\mathcal{Y}} : M(A) \geq \underline{R}(A), \forall A \in \mathcal{F}_{\mathcal{Y}}\}$ . Let  $\mathfrak{K}$  denote the class of all credal sets in  $\Delta_{\mathcal{Y}}$  and let

$$AU : \mathfrak{K} \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$$

*be an aleatoric uncertainty functional.*

*Define the feasible index set*

$$J_{AU} := \{j \in \{c, d\} : \mathcal{T}_j \in \mathfrak{K} \text{ and } AU(\mathcal{T}_j) \text{ is well-defined in } \overline{\mathbb{R}}\}.$$

*Then the optimization problem*

$$\min_{j \in \{c, d\}} AU(\mathcal{T}_j) \quad (9)$$

*has a solution (i.e.  $\arg \min_{j \in \{c, d\}} AU(\mathcal{T}_j) \neq \emptyset$ ) if and only if  $J_{AU} \neq \emptyset$ . Moreover, whenever  $J_{AU} \neq \emptyset$ ,*

$$\arg \min_{j \in \{c, d\}} AU(\mathcal{T}_j) = \begin{cases} \{c\}, & AU(\mathcal{T}_c) < AU(\mathcal{T}_d), \\ \{d\}, & AU(\mathcal{T}_d) < AU(\mathcal{T}_c), \\ \{c, d\}, & AU(\mathcal{T}_c) = AU(\mathcal{T}_d), \end{cases}$$

*(with the understanding that comparisons are taken in  $\overline{\mathbb{R}}$ ).*

If, in addition, one requires a finite optimal value, i.e.  $\min_{j \in \{c,d\}} AU(\mathcal{T}_j) \in \mathbb{R}$ , then (9) has a solution with finite value if and only if  $J_{AU} \neq \emptyset$  and

$$\min\{AU(\mathcal{T}_c), AU(\mathcal{T}_d)\} < +\infty.$$

*Proof.* By Proposition 2 (Conjunctive Fusion) and Proposition 1 (Disjunctive Fusion), under (8) both  $\mathcal{T}_c$  and  $\mathcal{T}_d$  are credal sets, hence belong to  $\mathfrak{K}$ . Therefore  $J_{AU} \neq \emptyset$  is equivalent to: at least one of the two values  $AU(\mathcal{T}_c)$ ,  $AU(\mathcal{T}_d)$  is well-defined in  $\overline{\mathbb{R}}$ .

( $\Rightarrow$ ) If (9) has a solution, then by definition there exists some  $j^* \in \{c,d\}$  for which  $AU(\mathcal{T}_{j^*})$  is well-defined and  $AU(\mathcal{T}_{j^*}) \leq AU(\mathcal{T}_j)$  for all  $j \in \{c,d\}$ . Hence  $j^* \in J_{AU}$  and  $J_{AU} \neq \emptyset$ .

( $\Leftarrow$ ) If  $J_{AU} \neq \emptyset$ , then the set of feasible objective values  $\{AU(\mathcal{T}_j) : j \in J_{AU}\}$  is a nonempty finite subset of the totally ordered set  $\overline{\mathbb{R}}$ , hence it admits a minimum. Any index  $j^*$  attaining that minimum solves (9). The explicit form of the argmin set follows by comparing the two numbers  $AU(\mathcal{T}_c)$  and  $AU(\mathcal{T}_d)$  in  $\overline{\mathbb{R}}$ .

The final statement about finiteness is immediate: the optimal value is finite if and only if at least one feasible value is finite, equivalently

$$\min\{AU(\mathcal{T}_c), AU(\mathcal{T}_d)\} < +\infty. \quad \square$$

**Remark 2** (Minimal hypotheses in our setup). *Under Assumption III-A,  $\mathcal{T}_c$  is always nonempty. Independently of Assumption III-A,  $\mathcal{T}_d$  is nonempty because it contains  $\mathcal{P} \cup \mathcal{Q}$ . Thus, in our setting the only remaining requirement for existence of a minimizer is that  $AU$  is well-defined (as an  $\overline{\mathbb{R}}$ -valued functional) on both fused credal sets.*

**Lemma 2** (Set inclusion between the two fusions). *Assume  $\mathcal{P}, \mathcal{Q} \subseteq \Delta_{\mathcal{Y}}$  are nonempty and let*

$$\mathcal{T}_c := \mathcal{P} \cap \mathcal{Q}, \quad \mathcal{T}_d := \mathcal{M}(\underline{R}), \quad \text{where } \underline{R} = \min\{\underline{P}, \underline{Q}\}.$$

Then

$$\mathcal{T}_c \subseteq \mathcal{T}_d.$$

*Proof.* Let  $M \in \mathcal{T}_c = \mathcal{P} \cap \mathcal{Q}$ . Then in particular  $M \in \mathcal{P}$ , hence  $M(A) \geq \inf_{P \in \mathcal{P}} P(A) = \underline{P}(A)$  for all  $A \in \mathcal{F}_{\mathcal{Y}}$ . Therefore

$$M(A) \geq \underline{P}(A) \geq \min\{\underline{P}(A), \underline{Q}(A)\} = \underline{R}(A)$$

for all  $A$ , so  $M \in \mathcal{M}(\underline{R}) = \mathcal{T}_d$ .  $\square$

**Definition 1** (Inclusion monotonicity / antitonicity). *Let  $\mathfrak{K}$  be the class of credal sets (nonempty, convex, setwise closed subsets of  $\Delta_{\mathcal{Y}}$ ). A functional  $AU : \mathfrak{K} \rightarrow \overline{\mathbb{R}}$  is called*

- inclusion-monotone (isotone) if  $K_1 \subseteq K_2 \Rightarrow AU(K_1) \leq AU(K_2)$ ;
- inclusion-antitone if  $K_1 \subseteq K_2 \Rightarrow AU(K_1) \geq AU(K_2)$ ;
- strictly inclusion-monotone if  $K_1 \subsetneq K_2 \Rightarrow AU(K_1) < AU(K_2)$ ;
- strictly inclusion-antitone if  $K_1 \subsetneq K_2 \Rightarrow AU(K_1) > AU(K_2)$ .

**Theorem 3** (Structural conditions ensuring which fusion minimizes AU). *Assume  $\mathcal{T}_c, \mathcal{T}_d \in \mathfrak{K}$  and  $AU(\mathcal{T}_c), AU(\mathcal{T}_d) \in \overline{\mathbb{R}}$  are well-defined. Then:*

(a) *If  $AU$  is inclusion-monotone, then  $\mathcal{T}_c$  is an AU-minimizer:*

$$AU(\mathcal{T}_c) \leq AU(\mathcal{T}_d).$$

*If  $AU$  is strictly inclusion-monotone and  $\mathcal{T}_c \subsetneq \mathcal{T}_d$ , then  $\mathcal{T}_c$  is the unique minimizer:*

$$AU(\mathcal{T}_c) < AU(\mathcal{T}_d).$$

(b) *If  $AU$  is inclusion-antitone, then  $\mathcal{T}_d$  is an AU-minimizer:*

$$AU(\mathcal{T}_d) \leq AU(\mathcal{T}_c).$$

*If  $AU$  is strictly inclusion-antitone and  $\mathcal{T}_c \subsetneq \mathcal{T}_d$ , then  $\mathcal{T}_d$  is the unique minimizer:*

$$AU(\mathcal{T}_d) < AU(\mathcal{T}_c).$$

*Proof.* By Lemma 2 we have  $\mathcal{T}_c \subseteq \mathcal{T}_d$ . If  $AU$  is inclusion-monotone, this implies  $AU(\mathcal{T}_c) \leq AU(\mathcal{T}_d)$ ; if  $AU$  is strictly inclusion-monotone and the inclusion is strict, the inequality is strict. The antitone cases are analogous.  $\square$

**Remark 3** (“Minimal” local versions). *The global assumptions in Definition 1 can be weakened to the local property on the specific pair:*

$$\mathcal{T}_c \subseteq \mathcal{T}_d \Rightarrow AU(\mathcal{T}_c) \leq AU(\mathcal{T}_d) \quad (\text{resp. } \geq),$$

*which is the weakest structural condition on  $AU$  that forces  $\mathcal{T}_c$  (resp.  $\mathcal{T}_d$ ) to be a minimizer.*

**Proposition 7** (Two generic AU templates). *Let  $a : \Delta_{\mathcal{Y}} \rightarrow \overline{\mathbb{R}}$  be any extended-real functional on single probabilities and define, for  $K \in \mathfrak{K}$ ,*

$$AU^{\text{sup}}(K) := \sup_{\mu \in K} a(\mu), \quad AU^{\text{inf}}(K) := \inf_{\mu \in K} a(\mu).$$

*Then  $AU^{\text{sup}}$  is inclusion-monotone and  $AU^{\text{inf}}$  is inclusion-antitone. Consequently, in the fusion problem,  $AU^{\text{sup}}$  always selects  $\mathcal{T}_c$  (weakly) while  $AU^{\text{inf}}$  always selects  $\mathcal{T}_d$  (weakly).*

*Proof.* If  $K_1 \subseteq K_2$ , then  $\sup_{\mu \in K_1} a(\mu) \leq \sup_{\mu \in K_2} a(\mu)$  and  $\inf_{\mu \in K_1} a(\mu) \geq \inf_{\mu \in K_2} a(\mu)$ , proving monotonicity/antitonicity. The fusion conclusion follows from Theorem 3.  $\square$

## V. EXPERIMENTS

### A. Experimental Objectives and Hypotheses

Our theoretical results in Theorem 2 and Theorem 3 establish that due to structural inclusion  $\mathcal{T}_c \subseteq \mathcal{T}_d$ , the optimal fusion operator is determined by the monotonicity of the chosen AU functional. In particular, inclusion-monotone AU favor conjunctive fusion, whereas inclusion-antitone AU favors disjunctive fusion. Beyond validating this structural prediction, our experiments aim to characterize the system-level behavior of the credal fusion under varying degrees of sensor disagreement. In practical multi-sensor systems, fusion operators implicitly encode a fusion design principle: conjunctive fusion is conservative and may become infeasible under strong conflict, while disjunctive fusion is permissive but may lead to overconfident predictions. We therefore evaluate whether AU-guided operator selection acts as a conflict-adaptive control mechanism that guides the fusion-choice in response to sensor

disagreement. We consider a controlled multi-sensor classification setting in which multiple sensor streams (homogeneous or heterogeneous) observe the same object. Each sensor is modeled by a neural network ensemble, inducing a credal set in the categorical probability simplex. Fusion is performed using either conjunctive or disjunctive rules, or via AU-based operator selection.

This leads to the following hypotheses:

- H1 Structural Consistency:** Minimizing an inclusion-monotone AU functional selects the conjunctive fusion operator, while minimizing an inclusion-antitone AU functional selects the disjunctive operator, in agreement with the theoretical inclusion structure.
- H2 Conflict Adaptivity:** The AU guided selection mechanism adapts the chosen fusion operator as sensor conflict increases, favoring conjunctive fusion in low-conflict regimes and disjunctive fusion in high-conflict regimes.
- H3 Practical Robustness:** AU-based operator selection improves or stabilizes predictive performance relative to fixed strategies that always apply conjunctive or always apply disjunctive fusion, particularly in moderate to high-sensor conflict regimes.

## B. Experimental Design

For the experiments, we consider two and three view settings. For the controlled augmentations for simulating different sensor streams, we use mixtures of clean and covariate-shifted inputs (e.g., Gaussian noise, color jitter). Our objective is to evaluate whether theoretically grounded fusion improves decision quality (measured in terms of accuracy, negative-log likelihood, and Brier score) under such shifts in the input.

*a) Datasets and Model Architectures:* We run our proof of concept experiments on CIFAR [14] benchmarks (CIFAR10 and 100) and SVHN [15] dataset. To induce the credal sets, we train ensembles of ResNet [38] variants (ResNet20, ResNet32, and ResNet56). Specifically, we either vary the seeds only to generate an ensemble of the same architecture or vary the architecture definition as well (e.g., for a five-member ensemble, we use 2 ResNet20, 2 ResNet32, and 1 ResNet56 architecture for CIFAR100). In our experiments, class scores are converted into categorical predictive distributions via softmax normalization, and the resulting ensemble family is used as an empirical approximation to a credal set. We clarify which type of ensemble is being used during the discussion of their respective experimental results. Training is performed on clean data with only random horizontal flip augmentation unless otherwise specified.

*b) Ensemble Training:* To train the ensembles, we use a recent robust ensemble training mechanism CreDRO [9]

*c) Fusion Operators and AU-Based Selection.:* For each input, we construct: the conjunctive fused credal set  $\mathcal{T}_c$  and the disjunctive fused credal set  $\mathcal{T}_d$ . Given an AU functional  $AU(\cdot)$ , the selected operator minimizes the AU functional value. We evaluate both  $AU^{\text{sup}}$  (inclusion-monotone) and  $AU^{\text{inf}}$  (inclusion-antitone) variants to test consistency with Theorem 3. While we report the results for entropy, we observe similar trend for other AU measures (e.g., gini).

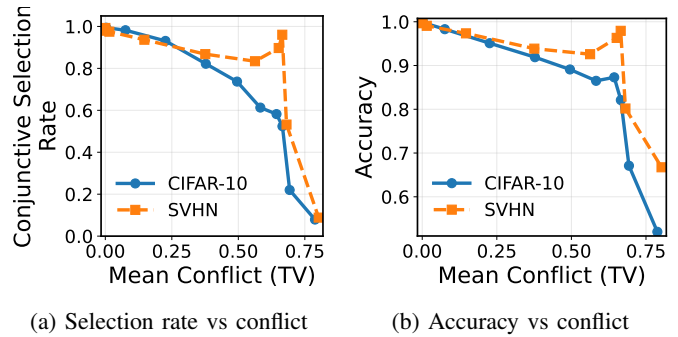


Fig. 2: Conflict-adaptive behavior of AU-based fusion. Conjunctive selection decreases monotonically with conflict, while AU-selection stabilizes predictive accuracy across regimes.

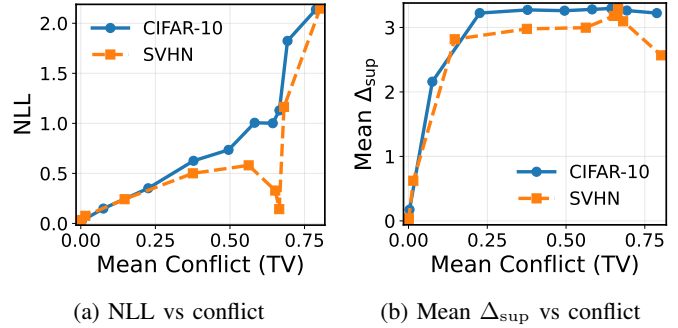


Fig. 3: Calibration and structural mechanism under conflict. AU-gap increases with disagreement.

## C. Evaluation Protocol

*a) Point Extraction from Credal Sets:* Credal fusion produces set-valued predictive outputs, while standard predictive metrics (e.g., accuracy or NLL) require a single probability distribution. We therefore extract a representative distribution from each fused credal set to enable decision-level evaluation. We use the **Barycenter** to obtain the predictive distribution where we take the mean of the fused credal set, used as the default representative distribution for all conjunctive, disjunctive, and AU-selected fusion strategies. This provides a neutral and consistent decision-level comparison across fusion operators.

*b) Predictive and Calibration Metrics:* We evaluate fused predictions using complementary performance criteria: **Accuracy** measures discrete decision correctness; **Negative Log Likelihood (NLL)** penalizes overconfident mispredictions and reflects decision-aligned risk; **Brier Score** evaluates squared probabilistic error; and **Expected Calibration Error (ECE)** quantifies probabilistic calibration.

*c) Conflict-Conditioned Evaluation:* Since the proposed method aims to adapt fusion behavior under sensor disagreement, we additionally analyze performance as a function of inter-sensor conflict. Conflict is quantified using total variation and Jensen–Shannon divergence between sensor barycenter predictions. We report conflict-binned performance curves and selection rates, allowing us to assess conflict-adaptive mechanism behavior of AU-guided fusion.

## D. Structural Consistency and Operator Selection (H1)

We first validate the structural consistency of AU-based operator selection predicted by Theorems 2 and 3. Figure 4

TABLE I: Structural consistency under  $AU^{\text{sup}}$  minimization (three-view setting). Violation rate measures empirical reversals of the theoretical ordering  $AU^{\text{sup}}(T_c) \leq AU^{\text{sup}}(T_d)$ .

| Dataset  | Samples | Sel. Conj | Viol. Rate (%) |
|----------|---------|-----------|----------------|
| CIFAR-10 | 10000   | 0.636     | 0.00           |
| SVHN     | 26032   | 0.806     | 0.00           |

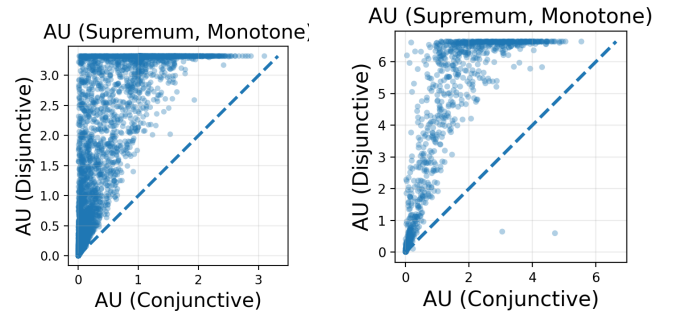
TABLE II: Structural consistency under  $AU^{\text{sup}}$  minimization (two-view setting). Violation rate measures empirical reversals of the theoretical ordering  $AU^{\text{sup}}(T_c) \leq AU^{\text{sup}}(T_d)$ .

| Dataset   | Samples | Sel. Conj | Viol. Rate (%) |
|-----------|---------|-----------|----------------|
| CIFAR-10  | 10000   | 0.991     | 0.02%          |
| CIFAR-100 | 10000   | 0.958     | 0.04%          |

illustrates the empirical relationship between  $AU^{\text{sup}}(T_c)$  and  $AU^{\text{sup}}(T_d)$  in the two-view setting. All samples lie above the diagonal, confirming the prediction from Theorem 3. Tables II and I quantify this structural consistency in both two-view and three-view configurations. In the two-view case (CIFAR-10 and CIFAR-100), violation rates remain below 0.05%, confirming near-perfect agreement with the theoretical ordering. In the more challenging three-view setting (CIFAR-10 and SVHN), the violation rate is identically zero across all samples. Importantly, while the conjunctive selection rate varies across datasets (e.g., 0.636 for CIFAR-10 and 0.806 for SVHN in the three-view setting), the violation rate remains zero. This distinction highlights that selection frequency reflects conflict and feasibility characteristics of the dataset, whereas violation rate measures consistency with the structural monotonicity property. In contrast, the inclusion-antitone functional  $AU^{\text{inf}}$  exhibits substantially higher empirical violation rates (e.g., over 60% in the three-view CIFAR-10 setting). These reversals do not contradict the order-theoretical property, but reflect the boundary-dominated nature of the infimum functional. Since  $AU^{\text{inf}}$  is determined by extremal vertices of the credal set, small numerical perturbations in high-conflict regimes can induce ordering instabilities. By comparison, the supremum-based functional  $AU^{\text{sup}}$  remains empirically stable across all datasets. Together, these results confirm Hypothesis H1: empirical operator selection aligns with the inclusion structure of credal fusion operators and the order-theoretic properties of the chosen AU functional.

### E. Conflict-adaptive operator selection (H2)

Figure 2 reports the fraction of samples selecting conjunctive fusion as a function of total-variation (TV) conflict. Across both CIFAR-10 and SVHN in the three-view setting, conjunctive fusion is selected almost exclusively in low-conflict regimes and decreases monotonically as conflict increases. In the highest-conflict bins, conjunctive selection drops substantially, indicating a transition toward disjunctive behavior. This smooth, monotonic trend demonstrates that *AU-based selection acts as a conflict-adaptive mechanism rather than enforcing a static fusion rule*. Importantly, the same qualitative behavior is observed across datasets despite differences in conflict distribution, confirming that operator



(a) CIFAR-10:  $AU^{\text{sup}}$

(b) CIFAR-100:  $AU^{\text{sup}}$

Fig. 4: Scatter plots illustrating the relationship between monotonic AU functionals. Under conjunctive fusion for CIFAR-10 and CIFAR-100, we observe that most of the inputs are above the diagonal line meaning that they favor conjunctive fusion. TABLE III: Global predictive performance in the three-view setting. Metrics are computed using barycenter extraction.

| Dataset  | Method      | Accuracy $\uparrow$ | NLL $\downarrow$ | Brier $\downarrow$ | ECE $\downarrow$ |
|----------|-------------|---------------------|------------------|--------------------|------------------|
| CIFAR-10 | Conjunctive | 0.683               | 0.904            | 0.347              | <b>0.031</b>     |
|          | Disjunctive | 0.833               | 1.947            | 0.767              | 0.622            |
|          | AU-selected | <b>0.849</b>        | 0.898            | 0.346              | 0.195            |
|          | Linear Pool | 0.833               | <b>0.524</b>     | <b>0.268</b>       | 0.130            |
| SVHN     | Conjunctive | 0.792               | 0.540            | 0.209              | <b>0.021</b>     |
|          | Disjunctive | 0.864               | 1.606            | 0.630              | 0.514            |
|          | AU-selected | <b>0.923</b>        | 0.525            | 0.204              | 0.122            |
|          | Linear Pool | 0.893               | <b>0.401</b>     | <b>0.191</b>       | 0.130            |

switching is driven by structural disagreement rather than dataset-specific effects. These results validate Hypothesis H2.

### F. Predictive Performance Under Fusion Strategies (H3)

Table III reports global predictive performance in the three-view setting. On CIFAR-10, conjunctive fusion is best calibrated (ECE = 0.031) but exhibits lower accuracy (0.683), reflecting its conservative behavior under disagreement. Disjunctive fusion improves accuracy (0.833) but suffers from substantially higher NLL and Brier scores, indicating overconfidence in high-conflict regimes. AU-guided selection achieves the highest accuracy among the credal operators (0.849) while maintaining NLL and Brier scores comparable to conjunctive fusion, thereby mitigating the brittleness of strict intersection without incurring the instability of disjunctive fusion. Linear pooling attains the lowest NLL and Brier score, consistent with its averaging behavior. A similar pattern is observed on SVHN. Conjunctive fusion remains well calibrated (ECE = 0.021) but less accurate (0.792), whereas disjunctive fusion improves accuracy (0.864) at the cost of degraded probabilistic metrics. AU-guided selection reaches 0.923 accuracy, substantially outperforming both fixed credal operators while preserving competitive calibration. Overall, across both datasets, AU-based operator selection consistently improves or stabilizes performance relative to fixed conjunctive and disjunctive strategies, validating Hypothesis H3, particularly in disagreement-heavy regimes.

### G. Statistical Stability

All results are averaged over five independent runs varying the random seeds and the observed standard deviations are on

the order of  $10^{-4}$  (less than 0.02% variation in accuracy). Hence, we omit them from the tables for better readability.

## VI. CONCLUSION

We presented a decision-driven framework for credal information fusion in multi-source systems. We showed that conjunctive and disjunctive fusion generate nested credal sets, and that this inclusion structure determines an optimal fusion operator through the order-theoretic properties of the chosen aleatoric uncertainty functional. Empirical results in multi-sensor classification under distribution shift confirm the predicted operator ordering and demonstrate improved calibration and stability relative to fixed fusion strategies. Overall, our work turns fusion operator choice from a heuristic modeling decision into a principled consequence of uncertainty geometry and downstream objectives. In the future, we will extend the work for a mix of clean and adversarial input streams and a more aggressive noise schedule to test the practical limit of the fusion operation.

## VII. ACKNOWLEDGMENT

This work has been supported by the NSF under grants CCF-2218845, ECCS-2229472, and ECCS-2329013; by AFOSR under grant FA9550-23-1-0261; by ONR under grant N00014-23-1-2221; and by DARPA under Cooperative Agreement D25AC00374-00.

## REFERENCES

- [1] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 6–23, 1997.
- [2] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: A review of the state-of-the-art," *Information Fusion*, vol. 14, no. 1, pp. 28–44, 2013.
- [3] S. J. Julier and J. K. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations," in *Proceedings of the 1997 American Control Conference*, pp. 2369–2373, 1997.
- [4] A. Kendall and Y. Gal, "What uncertainties do we need in Bayesian deep learning for computer vision?," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [5] E. Hüllermeier and W. Waegeman, "Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods," *Machine Learning*, vol. 110, pp. 457–506, 2021.
- [6] P. Walley, *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall, 1991.
- [7] M. C. M. Troffaes and G. de Cooman, *Lower Previsions*. John Wiley & Sons, 2014.
- [8] M. Caprio, S. Dutta, K. J. Jang, V. Lin, R. Ivanov, O. Sokolsky, and I. Lee, "Credal bayesian deep learning," arXiv preprint, 2023.
- [9] K. Wang, G. A. Faza, F. Cuzzolin, S. L. Chau, D. Moens, and H. Hallez, "Learning credal ensembles via distributionally robust optimization," *arXiv preprint arXiv:2602.08470*, 2026.
- [10] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press, 1976.
- [11] T. Denœux, "Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence," *Artificial Intelligence*, vol. 172, no. 2-3, pp. 234–264, 2008.
- [12] D. Dubois, W. Liu, J. Ma, and H. Prade, "The basic principles of uncertain information fusion: An organised review of merging rules in different representation frameworks," *Information Fusion*, vol. 32, pp. 12–39, 2016.
- [13] A. Bronevich and G. J. Klir, "Measures of uncertainty for imprecise probabilities: An axiomatic approach," *International Journal of Approximate Reasoning*, vol. 51, no. 4, pp. 365–390, 2010.
- [14] K. Alex, "Learning multiple layers of features from tiny images," <https://www.cs.toronto.edu/kriz/learning-features-2009-TR.pdf>, 2009.
- [15] I. J. Goodfellow, Y. Bulatov, J. Ibarz, S. Arnoud, and V. Shet, "Multi-digit number recognition from street view imagery using deep convolutional neural networks," *arXiv preprint arXiv:1312.6082*, 2013.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.
- [17] D. Dubois, W. Liu, J. Ma, and H. Prade, "The basic principles of uncertain information fusion. an organised review of merging rules in different representation frameworks," *Information Fusion*, vol. 32, pp. 12–39, 2016.
- [18] D. Dubois, W. Liu, J. Ma, and H. Prade, "Toward a general framework for information fusion," in *Modeling Decisions for Artificial Intelligence (MDAI 2013)*, vol. 8234 of *Lecture Notes in Computer Science*, pp. 37–48, Springer, 2013.
- [19] G. de Cooman and M. C. M. Troffaes, "Coherent lower previsions in systems modelling: products and aggregation rules," *Reliability Engineering & System Safety*, vol. 85, no. 1–3, pp. 113–134, 2004.
- [20] R. F. Nau, "The aggregation of imprecise probabilities," *Journal of Statistical Planning and Inference*, vol. 105, no. 2, pp. 265–282, 2002.
- [21] S. Bradley, "Aggregating belief models," in *Proceedings of the Eleventh International Symposium on Imprecise Probabilities: Theories and Applications*, vol. 103 of *Proceedings of Machine Learning Research*, pp. 38–48, PMLR, 2019.
- [22] K. Sentz and S. Ferson, "Combination of evidence in Dempster-Shafer theory," Tech. Rep. SAND2002-0835, Sandia National Laboratories, Albuquerque, NM, USA, 2002.
- [23] G. J. Klir, "Uncertainty and information measures for imprecise probabilities: An overview," in *Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications (ISIPTA '99)* (G. de Cooman, F. G. Cozman, S. Moral, and P. Walley, eds.), (Ghent, Belgium), pp. 234–240, 1999.
- [24] A. Bronevich and G. J. Klir, "Measures of uncertainty for imprecise probabilities: An axiomatic approach," *International Journal of Approximate Reasoning*, vol. 51, no. 4, pp. 365–390, 2010.
- [25] J. Abellán, G. J. Klir, and S. Moral, "Disaggregated total uncertainty measure for credal sets," *International Journal of General Systems*, vol. 35, no. 1, pp. 29–44, 2006.
- [26] E. Hüllermeier, S. Destercke, and M. H. Shaker, "Quantification of credal uncertainty in machine learning: A critical analysis and empirical comparison," in *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence*, vol. 180 of *Proceedings of Machine Learning Research*, pp. 548–557, PMLR, 2022.
- [27] Y. Sale, M. Caprio, and E. Hüllermeier, "Is the volume of a credal set a good measure for epistemic uncertainty?," in *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, vol. 216 of *Proceedings of Machine Learning Research*, pp. 1795–1804, PMLR, 2023.
- [28] F. Cuzzolin, *The geometry of uncertainty: the geometry of imprecise probabilities*. Springer Nature, 2020.
- [29] P. Hofman, Y. Sale, and E. Hüllermeier, "Quantifying aleatoric and epistemic uncertainty with proper scoring rules," 2024.
- [30] T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," *Journal of the American Statistical Association*, vol. 102, no. 477, pp. 359–378, 2007.
- [31] T. Seidenfeld, M. J. Schervish, and J. B. Kadane, "Forecasting with imprecise probabilities," *International Journal of Approximate Reasoning*, vol. 53, no. 8, pp. 1248–1261, 2012.
- [32] J. Konek, "Ip scoring rules: Foundations and applications," in *Proceedings of the Eleventh International Symposium on Imprecise Probabilities: Theories and Applications*, vol. 103 of *Proceedings of Machine Learning Research*, pp. 256–264, PMLR, 2019.
- [33] J. Konek, "Evaluating imprecise forecasts," in *Proceedings of the Thirteenth International Symposium on Imprecise Probability: Theories and Applications*, vol. 215 of *Proceedings of Machine Learning Research*, pp. 270–279, PMLR, 2023.
- [34] C. Fröhlich and R. C. Williamson, "Scoring rules and calibration for imprecise probabilities," *arXiv preprint arXiv:2410.23001*, 2024.
- [35] A. Singh, S. L. Chau, and K. Muandet, "Truthful elicitation of imprecise forecasts." <https://openreview.net/pdf?id=zuqYi3haGg>, 2025. OpenReview preprint.
- [36] I. Levi, *The Enterprise of Knowledge*. London, UK : MIT Press, 1980.
- [37] S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci, "Ergodic theorems for lower probabilities," *Proceedings of the American Mathematical Society*, vol. 144, no. 8, pp. 3381–3396, 2016.
- [38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.